

eleMEK

ELEMEK A MISKOLCI EGYETEMEN

Burmeister Erzsébet, erzsi@marki.lib.uni-miskolc.hu

Miskolci Egyetem, Könyvtár, Levéltár, Múzeum

Perlaki Attila, perlaki@kvtlinux.lib.uni-miskolc.hu

Miskolci Egyetem, Könyvtár, Levéltár, Múzeum

A metaadatok kezelése azok számára sem mindig világos, akik pedig, akár tudtukon kívül, naponta dolgoznak velük. Az eleMEK alkalmazás-csomag csak akkor bizonyul hasznos szerszámnak, ha tudjuk, mire használjuk, ezért kezdjük a metaadatoknál, a leíró rekordoknál.

A metaadatok az egyre növekvő információhalmazban való eligazodásunkat segítik, a kategorizálást, a keresést teszik könnyebbé. Ezek a könyvtárakban váltak tudománnyá, s alakultak ki igen összetett szerkezetekké, ám a mindennapi életben is számos esetben alkalmazzuk őket. Ahogy azonban egyre több az információ, a megszokott leírók szegényessé, használhatatlanná válnak.

Metaadat az is, amit egy operációs rendszer gépünkön őrzött állományainkról szolgáltat, ha listázási parancsot adunk ki. Operációs rendszertől függenek a részletek, de általában legalább két szintről kapunk információt:

- Megtudjuk, hogy melyik meghajtón, könyvtárban (directory, folder) vagyunk, ennek mi a neve, elérési útja, mekkora helyet foglal és esetleg azt is, mikor hoztuk létre.
- Megtudjuk a benne tárolt állományok neveit, méretüket, létrehozási dátumait, jogosultságait, esetleg egy kisebb képi ikont is – vagy közvetlenül, vagy egy-két kattintás után.

A lényeg tehát, hogy nem magát az adatot, hanem az adatról szóló adatokat kaptuk meg. Ez alkalmas korlátozott keresésekre (pl. keresem a 2004-es adóbevallásom tábláit), kategorizálásra (kiválogatom azokat a fotókat, amiket télen készítettem), de általában csak emberi segítséggel megy, automatikusan nem. Ennek az oka az, hogy a szükséges további metaadatok jó részét a fejünkben tároljuk még mindig (például tudom, hogy a digitális fényképezőgépem DSC betűhármassal kezdődő állományneveket állít elő, tehát a sok kép közül valószínű, bár nem biztos, hogy ezek általam készítették). Alaposan elgondolkodva bizonyára előnyösebb lenne, ha nem nekünk kellene megjegyezni, egy-egy állománynév, állománytípus, vagy egy-egy könyvtár mit is azonosít, hanem tetszőleges, vagy legalább is a mainál sokkal bővebb leírásokkal láthatnánk el anyagainkat.

Ha az állománynév nem mond semmit, ma csak az egyenkénti „belenézegetés” marad lehetőségnek. Ami pedig több tízezer állománynál már reménytelen feladatnak tűnik.

A több tízezer állomány a legutóbbi évekig olyan mennyiség volt, ami még kisvállalatokat sem mindig fenyegetett. Egy magánszemély csak gyűjtőként szembesült a katalogizálás problémáival. Egy (vállalati) kiskönyvtár néhány száz kötetével is „fejből” elboldogulhatott a sok év rutinjával rendelkező könyvtáros. Ennek azonban már vége, egyetlen digitális fényképezőgép is több száz fotót tud egy kártyán tárolni, amit másodpercek alatt lehet „átönteni” a számítógépre, majd onnan akár az internetre. Visszafelé pár nagyságrenddel még több állomány kerülhet a gépre. A „hova is tettem azt a képet, szöveget?” kérdés egyre nehezebben megválaszolható.

A könyvtárak már a digitális gyűjtemények kora előtt jóval ismerték és használták a metaadatokat, ha nem is éppen így hívták őket. Ha rátekintünk egy ún. katalógus cédulára, azon a legfontosabb információkat megtaláljuk, általában hamarabb, mintsem a polcok közti bolyongással – már amennyiben egyáltalán kint van a könyv és nem raktárban. Megtudjuk a szerzőt és a pontos címet, a kiadót, a kiadás dátumát, néhány azonosítókodeket, a témakört, néhány kulcsszót és a könyv megtalálását segítő információt. Ha publikációt keresünk, általában annak rövid összefoglaló leírását is olvashatjuk, mely segít dönteni, valóban ezt keressük-e? Ráadásul a katalóguscédula csak az egyik, tulajdonképpen a legegyszerűbb és legegyszerűbb metaadat a könyvtárak által használtak közül.

Talán érthető, hogy amikor a digitális gyűjtemények metaadatkezelésének problémáját megoldani kívánták, elsősorban a könyvtárak tapasztalataira építettek, ám ennek mutatkozott néhány árnyoldala is.

Az egyik, hogy a könyvtáraknál kialakult összetett leírási szabályok egyrészt túl bonyolultak voltak más irányú felhasználáshoz, másrészt sok esetben nem voltak megfelelően illeszthetők az adott problémákhoz. Alkalmas szabvány nem lévén a helyi jellegű megoldások szaporodtak el, az ezek közti adatcsere gyakorlatilag gazdaságosan kivitelezhetetlen volt. Hozzájárult még ehhez, hogy míg a könyvtárakban szakképzett személyek végezték a leírásokat, addig a digitális gyűjteményeknél elemi hibákkal és hiányosságokkal terhelt metaadat-bázisok születtek.

Az eleMEK, mint digitális gyűjteményekhez szánt eszköz, egy másik, már ismert és hasznosságát kiválóan bizonyított rendszer tapasztalataira épített. Ez a Magyar Elektronikus Könyvtár. Az eleMEK lényege nem csak az a szoftver, amely működteti, ez ugyanis lecserélhető. A kulcs az az adatszerkezet, amit a Drótos László által vezetett csapat kidolgozott és kvázi-szabványként elterjesztésre javasolt. Ebből az adatszerkezetből ugyanis ismert könyvtári szabványokra is lehetséges a konverzió, de ami még lényegesebb, lehetővé teszi, hogy az eleMEK-et használó gyűjtemények egymással könnyen adatot cseréljenek és közösen kereshetők legyenek. A rendszer pedig természetesen internet-alapú.

Be kell ismerni, hogy általában véve is az interneten elkésettnek látszik a metaadatkezelés szabványosított bevezetése. Túl sok leíratlan digitális dokumentum került már fel, melyek közt legfeljebb a teljes szövegű keresőkkel van valami remény megtalálni a keresett információt, rendezettségéről szó sincs. Az internet hőskorában még léteztek elképzelhetetlen munkaigénnyel létrehozott tematikus katalógusok, ezeket azonban, ha csak nem speciális célúak, az adattömeg szó szerint maga alá temette.

Az utólagos katalogizálás az erre váró anyag mennyiségétől függően vagy felesleges erőfeszítésnek tűnik (mert kevés az anyag, „úgy is tudom, mi hol van”), vagy reménytelennek („ki fog tizenkétezer képet átnézni és egyenként leírni?”). Mégis, valami ilyesmire szeretne az eleMEK rávenni, cserébe a későbbiekben egy kis rendszeres pluszmunkát igényelve az új anyagok befogadásakor és sok felhasználónak sok-sok idejét megtakarítva a kereséskor. Valamint esélyt arra, hogy megtalálja azt, amit keres.

Az a dokumentum, ami nincs meg, nem is létezik, hiába készült el.

A Miskolci Egyetemen azt tűztük ki célul, hogy a tanszékeken, intézetekben készülő publikációkat, oktatási segédleteket, a diákok által készített dolgozatokat és diplomamunkákat az ad hoc kialakított, vagy meg sem levő gyűjtési helyek helyett, vagy mellett, központilag is gyűjtsük. A létrehozandó digitális gyűjtemény metaadatkezelőjének pedig az eleMEK-et választottuk.

Az eleMEK tehát nem kölcsönzői rendszer, nem teljes értékű nyilvántartó és adatkezelő eszköz, csupán csak egy fontos segédeszköz, mely rugalmasan képes beépülni a helyére és a teljes rendszeren belül a metaadatok alapján való tájékozódást biztosítja az olvasóknak.

Az eleMEK egy Linux szerverre került fel, keresőszolgáltatása a könyvtári weblapok közé lett beépítve. Adatbázisának feltöltését itt szakképzett dolgozó végzi, azonban a rendszer működtetéséhez ez nem alapkövetelmény, a mellékelt igen részletes súgó alapján, ha nem is pár perc alatt, de elsajátíthatók a pontos leírás legalapvetőbb fogásai.

A választás azért is esett az eleMEK-re, mert beépülő eszközként nem kényszerít a webes oldal kialakításában a felhasználóra súlyos megkötéseket. Nyílt forrású moduláris eszközként pedig a továbbfejlesztése is lehetséges, egyedül az adatszerkezete kötött, érthető okokból.

Az egyik továbbfejlesztési terv az egyre sikeresebb Matarka rendszerrel való együttműködés. Itt az eleMEK a Matarka lekérdezési felülete és a teljes szöveg közé léphet be, gazdag metaadatleírási és szolgáltatási képességeivel.

A másik az EPA tapasztalataira épülő, periodikus kiadványok leírására szolgáló eleMEK-változat létrehozása. Ennek adatszerkezete már véglegesnek tekinthető, a programkód fejlesztése folyamatban van.

Célunk még annak elérése, hogy a metaadatleírás hasznosságát a dokumentum előállítóival is beláttassuk. Ez a tudás a keresésben is hasznos lehet. Talán az sem elhanyagolható apróság, hogy a rendszer által támogatott Creative Commons jogkezelést is szorgalmazzuk.

Hogy a szemantikus web végül kialakul-e és milyen formában, azt megjósolni, vagy befolyásolni az eleMEK valószínűleg nem fogja. De ha kialakul, használói előnnyel indulnak, számukra ez az új web nem lesz ismeretlen terület.

Irodalom:

Drótos László: eleMEK Metaadatkezelő rendszer digitális gyűjteményekhez

TMT 2005/2

<http://elemek.oszk.hu/ismertetok/elemek-cikk.htm>

Drótos László – Perlaki Attila: Moduláris könyvtári rendszer elektronikus dokumentumgyűjtemények kezeléséhez: az eleMEK projekt Networkshop, Győr, 2002. április

<http://elemek.oszk.hu/ismertetok/netwshp2004.htm>

Drótos László – Perlaki Attila: eleMEK rendszerterv folyamatosan aktualizált lap

<http://elemek.niif.hu/rendszerterv/index.html>

Drótos László – Góczán Andrea: eleMEK kitöltési útmutató folyamatosan aktualizált lap

<http://elemek.niif.hu/help/index.htm>

Burmeister Erzsébet – Kiss Andrea: Matarka dokumentumok folyamatosan aktualizált gyűjtőlap

<http://www.matarka.hu/other/dok.php>